

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2002, 08:20:15 ; Search time 22.96 Seconds

(Without alignments)
131.539 Million cell updates/sec

Title: US-09-351-778a-9

Sequence: 1 MTGSTIAPTDTYRTATATGL.....ICCKRRRAPPIYPIIVL 78

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	101	E311_ADE02	P24935 human adeno
2	38	48.7	101	E311_ADE06	O55653 human adeno
3	24	30.8	93	E311_ADE05	P17560 human adeno
4	6	7.7	92	HRK_MOUSE	P70678 mus musculu
5	6	7.7	121	VNRB_PORNY	O02123 poplar mosa
6	6	7.7	126	NIFM_FRASE	092570 frankia sp.
7	6	7.7	136	YRNB_CAEEL	O09420 caenorhabdi
8	6	7.7	142	YB9E_YEAST	P38337 saccharomyc
9	6	7.7	152	RK34_SPTOL	P82244 splinacia ol
10	6	7.7	157	HE52_MOUSE	O54792 mus musculu
11	6	7.7	157	HE52_RAT	P35429 rattus norv
12	6	7.7	173	HE52_HUMAN	O93543 homo sapien
13	6	7.7	195	NRTN_MOUSE	P97463 mus musculu
14	6	7.7	248	RIBB_RHOOP	O84990 rhodococcus
15	6	7.7	248	VGL2_EBV	P03218 Epstein-Bar
16	6	7.7	263	Y225_METUA	O57678 methanococ
17	6	7.7	280	HE51_HUMAN	O14469 homo sapien
18	6	7.7	281	HE51_RAT	O04666 rattus norv
19	6	7.7	282	HE51_MOUSE	P35428 mus musculu
20	6	7.7	290	HE51_CHICK	O57337 gallus gall
21	6	7.7	290	YCAF_YEAST	P42938 saccharomyc
22	6	7.7	318	FERB_ECOLI	P14609 escherichia
23	6	7.7	325	YH03_YEAST	P38844 saccharomyc
24	6	7.7	328	SYFA_HELPY	O92K19 helicobacte
25	6	7.7	328	SYFA_HELPY	P56146 helicobacte
26	6	7.7	337	HAIR_DROME	P14003 drosophila
27	6	7.7	350	OY08_SYNY3	O55880 synecocyst
28	6	7.7	361	OXDA_FUSSO	P24552 fusarium so
29	6	7.7	369	SPTL_MYXVL	P12383 myxoma viru
30	6	7.7	377	TTL_BOVIN	P38564 bos taurus
31	6	7.7	378	HAIR_DROYI	P29303 drosophila
32	6	7.7	379	TTL_PIG	P38160 sus scrofa
33	6	7.7	398	HEM1_ECOLI	P09128 escherichia

34	6	7.7	401	1	CSD_PSEPU	O92408 pseudomonas
35	6	7.7	408	1	AATC_CAEEL	O22067 caenorhabdi
36	6	7.7	410	1	NORB_PASML	O9C1B0 pasteurilla
37	6	7.7	411	1	NORB_HAEIN	O05011 haemophilus
38	6	7.7	420	1	MLIC_XENLA	P49219 xenopus lae
39	6	7.7	435	1	DPN_DROME	O26263 drosophila
40	6	7.7	435	1	FIYC_RHIME	P09820 rhizobium m
41	6	7.7	440	1	YDYN_HAEIN	P45079 haemophilus
42	6	7.7	452	1	FABH_HORVU	P23902 hordeum vul
43	6	7.7	455	1	COBQ_LACICA	O43989 acinetobact
44	6	7.7	472	1	C13B_MYCTU	O53765 mycobacteri
45	6	7.7	481	1	UCDH_CAEEL	O19905 caenorhabdi

ALIGNMENTS

```

RESULT 1
ID      E311_ADE02      STANDARD;      PRT;      101 AA.
AC      P24935;
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      01-DEC-1992 (Rel. 24, Last annotation update)
DE      Early E3A 11.6 kDa glycoprotein.
OS      Human adenovirus type 2.
OC      Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX      NCBI_TaxID=10515;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=81053687; PubMed=6253880;
RA      Herlase J., Coutois G., Gallibert F.;
RT      "Nucleotide sequence of the EcoRI D fragment of adenovirus 2 genome.";
RL      Nucleic Acids Res. 8:2173-2192(1980).
RN      [2]
RP      IDENTIFICATION OF PROTEIN.
RX      MEDLINE=93079877; PubMed=1448922;
RA      Scarla A., Tollefsen A.E., Saha S.K., Mold W.S.M.;
RT      "The E3-11.6k protein of adenovirus is an Asn-glycosylated integral
      membrane protein that localizes to the nuclear membrane.";
RL      Virology 191:743-753(1992).
CC      -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC      -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED.
CC      -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
      kDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
      between the Swiss Institute of Bioinformatics and the EMBL outstation -
      the European Bioinformatics Institute. There are no restrictions on its
      use by non-profit institutions as long as its content is in no way
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (See http://www.isb-sib.ch/announce/
      or send an email to license@isb-sib.ch).
CC      EMBL: J01917; AAA92222.1;
DR      EMBL; J01917; AAA92222.1;
KW      Early protein; Glycoprotein; Transmembrane.
FT      TRANSMEM 41..62
FT      CARBOHYD 14..14 N-LINKED (GLCNAC...?) (POTENTIAL).
SQ      SEQUENCE 101 AA; 11644 MW; FB89FCC6E921E84B CRC64;

```

```

Query Match      100.0%; Score 78; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 MTGSTIAPTDTYRTATATGLTSLNLPVHAFFVNAMASLDMMWMSIALMPFCLIMMLIC 60
DB      1 MTGSTIAPTDTYRTATATGLTSLNLPVHAFFVNAMASLDMMWMSIALMPFCLIMMLIC 60
OY      61 CKRRRRAPPIYPIIVL 78
DB      61 CKRRRRAPPIYPIIVL 78

```

```
RESULT 2
E311_ADE06 STANDARD; PRT; 101 AA.
ID E311_ADE06
AC O53653;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Early E3A 11.6 kDa glycoprotein.
OS Human adenovirus type 6.
OC Viruses; dsDNA viruses; no RNA stage; Adenoviridae; Mastadenovirus.
NCBI_TaxID=10534;
RN [1]
RP SEQUENCE FROM N.A.
RA Reichmann H., Schaefer H., Geisler B., Hausmann J., Ortman D.,
RA Bauer U., Flunker G., Seidel W.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
CC KDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y16037; CAA75991.1;
KW Early protein; Glycoprotein; Transmembrane.
FT TRANSMEM 41
FT CARBOHYD 14
FT CARBOHYD 14
SO SEQUENCE 101 AA; 11613 MW; EF219000939E3B4B CRC64;

Query Match 48.78; Score 38; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.2e-34;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 MMFSTALMFVCLIMLIMLICKRRRARRPIRPIIVL 78
Db 41 MMFSTALMFVCLIMLIMLICKRRRARRPIRPIIVL 78

RESULT 3
E311_ADE05 STANDARD; PRT; 93 AA.
ID E311_ADE05
AC P17590;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Early E3A 10.5 kDa glycoprotein.
OS Human adenovirus type 5.
OC Viruses; dsDNA viruses; no RNA stage; Adenoviridae; Mastadenovirus.
NCBI_TaxID=28285;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-85092388; PubMed-2981456;
RA Claderes C., Wold W.S.M.;
RL "DNA sequence of the early E3 transcription unit of adenovirus 5.";
RL Virology 140:28-43(1985).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE-92087470; PubMed-1727603;
RX Chroboczek J., Biebert F., Jacrot B.;
RT "The sequence of the genome of adenovirus type 5 and its comparison
RT with the genome of adenovirus type 2.";
RL Virology 186:280-285(1992).
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED.
```

```
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
CC KDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M73260; NOT_ANNOTATED_CDS.
DR EMBL: X03002; CAA26784.1;
DR PIR: A05245; ERA053.
KW Early protein; Glycoprotein; Transmembrane.
FT TRANSMEM 34
FT CARBOHYD 3
FT CARBOHYD 3
SO SEQUENCE 93 AA; 10523 MW; 008AD087AB45A8F CRC64;

Query Match 30.88; Score 24; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 4.7e-19;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 MMFSTALMFVCLIMLIMLICKR 64
Db 34 MMFSTALMFVCLIMLIMLICKR 57

RESULT 4
HRK_MOUSE STANDARD; PRT; 92 AA.
ID HRK_MOUSE
AC P70678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE Activator of apoptosis harakiri (Neuronal death protein DPS).
GN HRK OR DPS OR BID3.
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Imatsumi K.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES APOPTOSIS AND INTERACTS SELECTIVELY WITH
CC SURVIVAL-PROMOTING PROTEINS BCL-2 AND BCL-XL (BY SIMILARITY).
CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR THE INDUCTION OF CELL
CC DEATH (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D83697; BA012065.1;
DR EMBL: D83698; BA012066.1;
DR MGD: MGI:1201608; BID3.
DR InterPro: IPR000712; BCL-2.
DR PROSITE: PS01259; BH3; 1.
KW Apoptosis.
FT DOMAIN 33
FT DOMAIN 47
SO SEQUENCE 92 AA; 10078 MW; 95760B9EAB88FIDE CRC64;

Query Match 7.78; Score 6; DB 1; Length 92;
```

Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 RRRARP 69
| | | | |
DB 50 RRRARP 55

RESULT 5

VNBP_POPMV STANDARD; PRT; 121 AA.

AC 002123;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 14 kDa protein (putative nucleic acid-binding protein).
OS Poplar mosaic virus (isolate ATCC PV275) (PMV).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX NCBI_TaxID=31709;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9233281; PubMed=1629709;
RA Henderson J., Gibbs M.J., Edwards M.L., Clarke V.A., Gardner K.A.,
RA Cooper J.I.,
RT Partial nucleotide sequence of poplar mosaic virus RNA confirms its
RT classification as a carlavirus.
RL J. Gen. Virol. 73:187-189(1992).
CC -1- FUNCTION: MAY ACT AS A REGULATORY FACTOR DURING VIRAL
CC TRANSCRIPTION.
CC -1- SIMILARITY: TO NUCLEIC ACID-BINDING PROTEINS FROM OTHER
CC CARLAVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X65102; CAA6227.1; -;
DR EMBL: D13364; BAA0229.1; -;
DR PIR: J01646; J01646;
DR PIR: S23876; S23876;
DR InterPro: IPR002568; Carla_C4.
DR Pfam: PF01623; Carla_C4; 1.
KW Zinc-finger; DNA-binding.
FT ZN-FING 62 83
FT C4-TYPE (POTENTIAL).
SQ SEQUENCE 121 AA; 14451 MW; 705205527BFD3CF CRC64;

Query Match 7.7%; Score 6; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 RRRARP 68
| | | | |
DB 52 RRRARP 57

RESULT 6

NIFW_FRASE STANDARD; PRT; 126 AA.

AC 0925Y0;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nitrogenase stabilizing/protective protein nifW.
GN NIFW.
OS Frankia sp. (strain EuIK1).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Frankineae; Frankiaceae; Frankia.
OX NCBI_TaxID=47227;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EuIK1;
RA Chung-Sun A., Ji-Tae K., Won-Jin K., Won-Young Y.;
RT "Nif-gene organization and nucleotide sequences from Frankia EuIK1
RT strain";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May protect the nitrogenase Fe-Mo protein from oxidative
CC damage (by similarity).
CC -1- SUBUNIT: Homotrimer; associates with nifD (by similarity).
CC -1- SIMILARITY: BELONGS TO THE NIFW FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF119361; AAD17267.1; -;
DR Nitrogen fixation.
KW SEQUENCE 126 AA; 14287 MW; 295D8C9102FD317E CRC64;

Query Match 7.7%; Score 6; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 TTATGL 20
| | | | |
DB 80 TTATGL 85

RESULT 7

YRN6_CAEEL STANDARD; PRT; 136 AA.

AC 009420;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Hypothetical 14.2 kDa protein R07B1.6 in chromosome X.
GN R07B1.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Kershaw J.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z48621; CAA8543.1; -;
DR WormPep; R07B1.6; CE01632.
DR InterPro: IPR002601; C6.
DR Pfam: PF01681; C6; 1.
KW Hypothetical protein.
SQ SEQUENCE 136 AA; 14169 MW; D7F47A83F3F81787 CRC64;

Query Match 7.7%; Score 6; DB 1; Length 136;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 T1AFT 10
 DB 31 T1AFT 36

RESULT 8

YB9E_YEAST STANDARD: PRT; 142 AA.

AC P38337;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical 16.4 kDa protein in POP4-SHM1 intergenic region.
 GN YBR238C OR YBR1726.

OS Saccharomyces cerevisiae (Baker's yeast).
 OC Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;

RN SEQUENCE FROM N.A.

RC STRAIN=5288C;

RX MEDLINE=93220397; PubMed=8465606;

RA Daignon F., Bileau N., Crouzet M., Aigle M.;

RT "The complete sequence of a 19,482 bp segment located on the right
 arm of chromosome II from Saccharomyces cerevisiae.";

RL Yeast 9:188-199(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

DR EMBL: X70529; CAA49822.1; -;
 DR EMBL: Z36127; CAA85221.1; -;
 DR PIR: S32960; S32960.
 DR SGD: S0000462; YBR238C.
 KM Hypothetical protein.
 SQ SEQUENCE 142 AA; 16429 MW; 12E73F0C3DD225D7 CRC64;

Query Match 7.7%; Score 6; DB 1; Length 142;

Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GLTSAL 24
 DB 71 GLTSAL 76

RESULT 9

RK34_SPTOL STANDARD: PRT; 152 AA.

ID RK34_SPTOL

AC P82244;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 50S ribosomal protein L34, chloroplast precursor.

GN RL34.

OS Spinacia oleracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllales; Caryophyllales; Chenopodiales; Spinacia.

OX NCBI_TaxID=3562;

RN SEQUENCE FROM N.A., SEQUENCE OF 92-101, AND MASS SPECTROMETRY.

RC STRAIN=CV. ALVARO; TISSUE=leaf;

RX MEDLINE=20435798; PubMed=10874046;

RA Yamaguchi K., Subramanian A.R.;

RT "The plastid ribosomal proteins. Identification of all the proteins in
 the 50 S subunit of an organelle ribosome (chloroplast).";

RL J. Biol. Chem. 275:28466-28482(2000).

CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.

CC -1- SUBCELLULAR LOCATION: Chloroplast.

CC -1- TISSUE SPECIFICITY: UBQUITOUS.

CC -1- MASS SPECTROMETRY: MW=6767.1; METHOD=Electrospray; RANGE=92-152.

CC -1- SIMILARITY: BELONGS TO THE L34P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

DR EMBL: AF238221; AAP6157.1; -;
 DR InterPro: IPR000271; Ribosomal_L34.

DR Pfam: PF00468; Ribosomal_L34; 1.

DR PROSITE: PS00784; RIBOSOMAL_L34; FALSE_NEG.

KW Ribosomal protein; Chloroplast; transit peptide; rRNA-binding.

FT TRANSIT 1 91

FT CHAIN 92 152

FT SEQUENCE 152 AA; 16095 MW; 755A90D441ADB18 CRC64;

Query Match 7.7%; Score 6; DB 1; Length 152;

Best Local Similarity 100.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 LKRRRA 67
 DB 127 LKRRRA 132

RESULT 10

HES2_MOUSE STANDARD: PRT; 157 AA.

ID HES2_MOUSE

AC 054792;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Transcription factor HES-2 (Hair cell enhancer of split 2).

GN HES2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129/J;

RX MEDLINE=98234545; PubMed=9570950;

RA Nishimura M., Isaka F., Ishidashi M., Tomita K., Tsuda H.,

RA Nakamishi S., Kageyama R.;

RT "Structure, chromosomal locus, and promoter of mouse Hes2 gene, a
 homologue of Drosophila hairy and enhancer of split.";

RL Genomics 49:69-75(1998).

CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
 PROTEIN FOR THEIR TRANSCRIPTION.

CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
 WITH A CO-REPRESSOR PROTEIN (GROUCHO).

CC -1- SUBCELLULAR LOCATION: Nucleus.

CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
 HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CAGNNG),

CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
 REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A

CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
 HAIRY-RELATED PROTEINS.

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AB009967; BAA24091.1; -
 DR MGP; MGI:1098624; Hes2.
 DR InterPro: IPR003015; HLH_MYC.
 DR InterPro: IPR001092; HLH_dlm.
 DR InterPro: IPR003650; Orange.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR SMART: SM00511; ORANGE; 1.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; 1.
 DR Transcription regulation; DNA-binding; Nuclear protein; Repressor.
 FT DNABIND 13 26
 FT DOMAIN 27 71 BASIC DOMAIN.
 FT DOMAIN 139 148 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 154 157 PRO-RICH.
 FT DOMAIN 154 157 WRPW MOTIF (REQUIRED FOR ACTIVITY)
 FT (BY SIMILARITY).
 SQ SEQUENCE 157 AA; 17231 MW; 570A0C67F4992EA7 CRC64;

Query Match 7.7%; Score 6; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 63 KRRRAR 68
 |||||
 Db 23 KRRRAR 28

RESULT 11
 HES2_RAT STANDARD; PRT; 157 AA.
 ID HES2_RAT
 AC P35429;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Transcription factor HES-2 (Hairly and enhancer of split 2).
 GN HES2 OR HES-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic brain;
 RX MEDLINE=9335886; Pubmed=8334270;
 RA Ishibashi M., Sasai Y., Nakanishi S., Kageyama R.;
 RT Molecular characterization of HES-2, a mammalian helix-loop-helix
 RT factor structurally related to Drosophila hairy and Enhancer of
 RT split.
 RT Eur. J. Biochem. 215:645-652(1993).
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
 CC PROTEIN FOR THEIR TRANSCRIPTION.
 CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
 CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
 CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG),
 CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
 CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
 CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO, A
 CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
 CC HAIRY-RELATED PROTEINS.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: D14029; BAA03118.1; -
 DR PIR: S35037; S35037.
 DR TRANSFAC: T01650.
 DR InterPro: IPR003015; HLH_MYC.
 DR InterPro: IPR001092; HLH_dlm.
 DR InterPro: IPR003650; Orange.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR SMART: SM00511; ORANGE; 1.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; 1.
 DR Transcription regulation; DNA-binding; Nuclear protein; Repressor.
 FT DNABIND 13 26
 FT DOMAIN 27 71 BASIC DOMAIN.
 FT DOMAIN 139 148 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 154 157 PRO-RICH.
 FT DOMAIN 154 157 WRPW MOTIF (REQUIRED FOR ACTIVITY)
 FT (BY SIMILARITY).
 SQ SEQUENCE 157 AA; 17028 MW; B5D621E814AE0369 CRC64;

Query Match 7.7%; Score 6; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 63 KRRRAR 68
 |||||
 Db 23 KRRRAR 28

RESULT 12
 HES2_HUMAN STANDARD; PRT; 173 AA.
 ID HES2_HUMAN
 AC Q9Y543; Q9Y542;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor HES-2 (Hairly and enhancer of split 2).
 GN HES2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Howden P.;
 RT Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
 CC PROTEIN FOR THEIR TRANSCRIPTION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
 CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG),
 CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
 CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
 CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO, A
 CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
 CC HAIRY-RELATED PROTEINS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AL031848; CAB46198.1; -
 DR EMBL: AL031848; CAB46199.1; -
 DR InterPro: IPR003015; HLH_Myc.
 DR InterPro: IPR001092; HLH_dlm.
 DR InterPro: IPR003650; Orange.
 DR Pfam: PF00010; HLH. 1.
 DR SMART: SM00353; HLH. 1.
 DR SMART: SM00511; ORANGE. 1.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX. 1.
 DR Transcription regulation; DNA-binding; Nuclear protein; Repressor;
 KW Alternative splicing.
 FT DNA_BIND 13 26 BASIC DOMAIN.
 FT DOMAIN 27 71 HELIX_LOOP_HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 139 167 PRO-RICH.
 FT DOMAIN 170 173 WRW MOTIF (REQUIRED FOR ACTIVITY)
 FT VARSPIC 48 173 NSNSKLEKADYLENTVFFLOELPSSWPIAPLPDSTRE
 FT GTSACVRLARVLPACRVLEPAVASARLLEHMRASATLD
 FT GGRAGDSGSPAPAPAPASAPASAPVSPSPGGLW
 FT RFW -> DASGWTWLPPLHAQNCPLTYIAPEQPPA (IN
 ISOFORM 2).
 SO SEQUENCE 173 AA; 18470 MW; B3751F4576B4F3D CRC64;

Query Match 7.7%; Score 6; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 KRRRAR 68
 111111
 DB 23 KRRRAR 28

RESULT 13
 NRTN_MOUSE STANDARD: PRT; 195 AA.
 ID NRTN_MOUSE
 AC P97463;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurturin precursor.
 GN NRTN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 AND
 RP 181-190.
 RX MEDLINE=97100947; PubMed=8945474;
 RA Kotbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
 RA Crendon D.J., Johnson E.M., Jr., Milbrandt J.,
 RT "Neurturin, a relative of glial cell-line-derived neurotrophic
 factor,"
 RL Nature 384:467-470(1996).
 CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
 CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
 CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
 CC HAEMOPOIETIC CELLS.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: U78109; AAC52954.1; -

DR HSSP: 007731; IAGQ.
 DR MGD; MGI:108417; Ntn.
 DR InterPro: IPR002400; GF_cysknoc.
 DR InterPro: IPR001839; TGF-beta.
 DR Pfam: PF00019; TGF-beta. 1.
 DR PRINTS: PR00438; GFCYSKNOC.
 DR SMART: SM00204; TGFBR. 1.
 DR PROSITE: PS00230; TGF-BETA_1; FALSE_NEG.
 DR Growth factor; Signal.
 KW SIGNAL 1 19
 FT PROPEP 95 95 POTENTIAL.
 FT CHAIN 20 95 BY SIMILARITY.
 FT DISULFD 96 195 NEURTURIN.
 FT DISULFD 101 163 BY SIMILARITY.
 FT DISULFD 128 192 BY SIMILARITY.
 FT DISULFD 132 194 BY SIMILARITY.
 FT DISULFD 162 162 INTERCHAIN (BY SIMILARITY).
 SO SEQUENCE 195 AA; 22219 MW; ABE21BB3D417448 CRC64;

Query Match 7.7%; Score 6; DB 1; Length 195;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 RRRARP 69
 111111
 DB 91 RRRARP 96

RESULT 14
 RIBF_RHOOP STANDARD: PRT; 248 AA.
 ID RIBF_RHOOP
 AC 084350;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative riboflavin biosynthesis protein ribF [includes: Riboflavin
 DE kinase (EC 2.7.1.26) (Flavokinase); FMN adenylyltransferase
 DE (EC 2.7.7.2) (FAD pyrophosphorylase) (FAD synthetase)].
 GN RIBF.
 OS Rhodococcus opacus (Nocardia opaca).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
 OC NCBI_TaxID=37919;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICP.
 RX MEDLINE=98324954; PubMed=9657989;
 RA Seibert V., Koubatova E.M., Golovleva L.A., Schloemann M.;
 RT "Characterization of a maleylacetate reductase encoding region from
 RT Rhodococcus opacus ICP."
 RL J. Bacteriol. 180:3503-3508(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + Riboflavin -> ADP + FMN.
 CC -1- CATALYTIC ACTIVITY: ATP + FMN -> diphosphate + FAD.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF030176; AAC38800.1; -
 DR Hypothetical protein; Transferase; Nucleotidyltransferase; FMN; FAD;
 KW Multifunctional enzyme.
 SO SEQUENCE 248 AA; 26703 MW; 7869F5B3D24AB56A CRC64;

Query Match 7.7%; Score 6; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 TTATGL 20

Db 7 TTTATG 12

```

RESULT 15
ID VGL2_EBV STANDARD; PRT; 248 AA.
AC P03218;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Probable membrane glycoprotein.
GN BILF2.
OS Epstein-Barr virus (strain 95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85035713; PubMed=6092825;
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RT "Sequence analysis of the 17,166 base-pair EcoRI fragment C of 95-8
RL Epstein-Barr virus."
RL Mol. Biol. Med. 1:21-45(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the 95-8 Epstein-Barr virus genome."
RL Nature 310:207-211(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90244363; PubMed=2159529;
RA Mackett M., Conway M.O., Arrand J.R., Haddad R.S.,
RA Hult-Fletcher L.M.;
RT "Characterization and expression of a glycoprotein encoded by the
RL Epstein-Barr virus BamHI I fragment."
RL J. Virol. 64:2545-2552(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V01555; CAA24803.1; -
CC DR EMBL: M37129; AAA45876.1; -
CC DR PIR: A03780; OQBEAL.
CC DR PIR: S33052; S33052.
CC DR InterPro: IPR003599; Ig.
CC DR InterPro: IPR003006; Ig_MHC.
CC DR Pfam: PF00047; Ig; 1.
CC DR SMART: SM00409; Ig; 1.
CC KW Glycoprotein; Late protein; Membrane.
SO SEQUENCE 248 AA; 27076 MW; C3F3A253B959ADA CRC64;

```

```

Query Match 7.7% Score 6; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 TTAPTT 10
Db 160 TTAPTT 165

```

Search completed: June 21, 2002, 08:25:25
Job time: 310 sec

